

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 06-165173

(43)Date of publication of application : 10.06.1994

(51)Int.Cl.

H04N 7/15  
H04M 3/56  
H04M 11/06  
H04N 7/18

(21)Application number : 04-307371

(71)Applicant : NIPPON TELEGR & TELEPH  
CORP <NTT>

(22)Date of filing : 17.11.1992

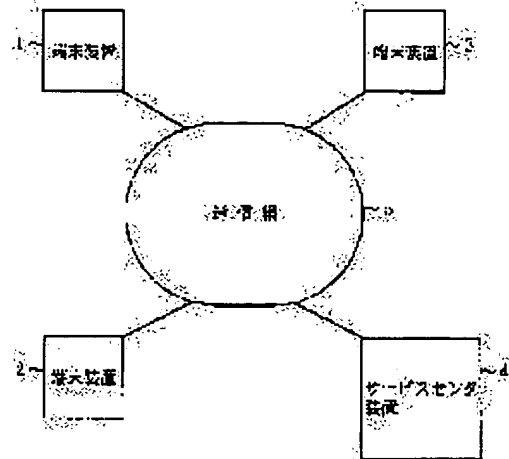
(72)Inventor : HIBINO YASUSHI  
SUZUKI HAJIME

## (54) SYSTEM FOR ATTAINING VIRTUAL SOCIETY

### (57)Abstract:

PURPOSE: To attain a virtual society as if many terminal users on remote places meet on the same place and make exchange.

CONSTITUTION: Each of terminal equipments 1 to 3 is provided with a photographing device, an image receiver, a transceiver and a control information input device. A service center equipment 4 is provided with an image synthesizer, a sound mixer, an image distributor, a sound distributor, a background image source, a background sound source, and a control device. Each terminal equipment sends user's image, sound and control information to a communication network 5. The equipment 4 synthesizes the images and sounds of respective users with the background image and sound to creat a virtual society. The image and sound of the synthesized virtual society are sent to the terminal equipments 1 to 3 for respective user's through the network 5.



## LEGAL STATUS

[Date of request for examination]

16.02.1999

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number] 3438737

[Date of registration] 13.06.2003

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

## \* NOTICES \*

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

---

CLAIMS

---

[Claim(s)]

[Claim 1] It consists of two or more terminal units and service center equipment connected with them through the communication line. Each terminal unit The image pick-up equipment which picturizes the person image of the terminal user concerned, and the telephone transmitter which inputs this user's voice, The input unit which inputs the control information by this user, and the sending set which transmits said user's person video signal, a sound signal, and control information to center equipment, The receiving set which receives the synthetic video signal which said center equipment transmits, and a mixed sound signal, It has the television equipment which outputs said received synthetic video signal, and the telephone transmitter which outputs a mixed sound signal. Service center equipment The source of a background image which accumulates two or more background video signals, and the background sound source which accumulates two or more background sound signals, The receiving set which receives the person video signal, the sound signal, and control information which said each terminal unit transmits, The image synthesizer unit which compounds the person video signal from said each terminal unit, and the background video signal chosen from said source of a background image, The voice mixing equipment which mixes the sound signal from said each terminal unit, and the background sound signal chosen from said background sound source, The distribution apparatus which distributes said synthetic video signal and mixed sound signal to each terminal unit, The virtual society implementation system characterized by what the control information from said each terminal unit is analyzed, and it has for the control unit which performs image composition of the selection control of the background video signal from said source of a background image, and a background sound source, or a background sound signal, said image synthesizer unit, and voice mixing equipment, and control of voice mixing.

[Claim 2] Said control unit is a virtual society implementation system according to claim 1 characterized by ordering said image synthesizer unit so that it may change to the object people video signal which the terminal unit concerned transmits as a synthetic video signal over the terminal unit concerned and the person silhouette video signal may be transmitted, when silhouette image directions are received from a certain terminal unit as control information.

[Claim 3] Said control unit is claim 1 characterized by ordering said image synthesizer unit so that the location of the person image of the terminal unit concerned in the inside of a background image may be made to move, or a virtual society implementation system given in two, when migration directions of the person image concerned are received from a certain terminal unit as control information.

[Claim 4] Said control unit is a virtual society implementation system according to claim 3 characterized by ordering said voice mixing equipment to make the sound volume of sound signals other than the terminal unit concerned adjust according to the location of the person image of the terminal unit concerned.

[Claim 5] Said control unit is a virtual society implementation system according to claim 3 which expands the person image of a specific terminal unit, makes voice level high with migration of the person image in the inside of a background image, and is characterized by ordering said image synthesizer unit and voice mixing equipment to reduce the person image from other terminal units, and

to cheat out of voice level low.

[Claim 5] Said control unit is a virtual society implementation system according to claim 1 or 2 which makes assignment selection of the visual field from the terminal unit concerned, expands the image within this visual field according to the control information from a certain terminal unit, makes voice level high, and is characterized by ordering said image synthesizer unit and voice mixing equipment to reduce the image besides a visual field and to cheat out of voice level low.

---

[Translation done.]

\* NOTICES \*

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

---

DETAILED DESCRIPTION

---

[Detailed Description of the Invention]

[0001]

[Industrial Application] This invention relates to the virtual society implementation system which provides two or more human beings with the place of an alternating current of imagination with a pictorial communication technique and a voice communication technique through a communication network.

[0002]

[Description of the Prior Art] The teleconference with voice is realized as a conventional similar system. Moreover, using the pictorial communication technique, at the so-called television conference, among two or more points, an image is sent, it suits, the image of a television machine is divided, and the method which displays two or more screens there at coincidence, and the method displayed on two or more television machines of airraid correspondence are taken.

[0003]

[Problem(s) to be Solved by the Invention] In the above-mentioned conventional technique, an alternating current in the same location as the partner who wants to alternate mutually cannot be performed, and there is a problem that the number of points is restricted, by communication between many points.

[0004] As the purpose of this invention is whether two or more alternating human beings gathered in one location, it is to offer the virtual society implementation system with an eye on losing human being's number of limits which the alternating current which led can be performed and interchange conversation and a gesture.

[0005]

[Means for Solving the Problem] The virtual society implementation system of this invention consists of two or more terminal units and service center equipment connected with them through the communication line. Each terminal unit The image pick-up equipment which picturizes the person image of the terminal user concerned, and the telephone transmitter which inputs this user's voice, The input unit which inputs the control information by this user, and the sending set which transmits said user's person video signal, a sound signal, and control information to center equipment, The receiving set which receives the synthetic video signal which said center equipment transmits, and a mixed sound signal, It has the television equipment which outputs said received synthetic video signal, and the telephone transmitter which outputs a mixed sound signal. Service center equipment The source of a background image which accumulates two or more background video signals, and the background sound source which accumulates two or more background sound signals, The receiving set which receives the person video signal, the sound signal, and control information which said each terminal unit transmits, The image synthesizer unit which compounds the person video signal from said each terminal unit, and the background video signal chosen from said source of a background video signal, The voice mixing equipment which mixes the sound signal from said each terminal unit, and the background sound signal chosen from said background sound source, The distribution apparatus which distributes said synthetic

video signal and mixed sound signal to each terminal unit, The control information from said each terminal unit is analyzed, and it is characterized by having the control unit which performs image composition of the selection control of the background video signal from said source of a background image, and a background sound source, or a background sound signal, said image synthesizer unit, and voice mixing equipment, and control of voice mixing.

[0006]

[Function] A user's image and voice which are in front of each terminal unit are changed into a video signal and a sound signal by image pick-up equipment and the telephone transmitter, and are sent out to a communication network by them, respectively. each of this user's image and sound signal are sent to service center equipment, are compounded with a background image and a background sound there, and are \*\*\*\* about the imagination world "virtual society" -- \*\*. Through a communication network, this image and voice of "virtual society" that were compounded are sent to each user's terminal unit, and are reproduced by television equipment and the earphone.

[0007]

[Example] Hereafter, a drawing explains one example of this invention.

[0008] Drawing 1 is the conceptual diagram of the whole virtual society implementation system of this invention, and terminal units 1, 2, and 3 being connected with service center equipment 4 through a communication network 5, and receiving service of virtual society is shown. In this example, although the example of three sets of terminals is shown, this invention does not limit the number of terminals.

[0009] Drawing 2 is the example of a configuration of each terminal units 1, 2, and 3, and consists of control information input units 15, such as the earphones 14, such as the television equipments 13, such as the telephone transmitters 12, such as the image pick-up equipments 11, such as a television camera, and a microphone, and a display, and a loudspeaker, a keyboard, and a mouse, and a transmitter-receiver 16 containing a sign / decryption section 161.

[0010] Here, the control information input unit 15 is used for the input of the following control information.

(1) It uses for selection of "virtual society." Thereby, a user can choose favorite "society" out of two or more "virtual society" prepared for service center equipment. | ✕

(2) Use in order to choose on which location of the space which "virtual society", i.e., a background image, and a background sound make a user image is put. By this, a user can move about in "virtual society." | ✕

(3) When returning a user image to the user concerned, use for choosing whether it considers as an image as it is, or it considers as a silhouette image.

(4) Use for directions to expand a specific partner's image (image of a rise), enlarge [ voice / of a partner ] alternatively, and for there be small other voice. Conversation with a specific partner is attained by this.

[0011] Drawing 3 is the example of a configuration of service center equipment 4. The transmitter-receiver 41 containing the sign / decryption sections 411, 412, and 413 for multiple channels (this example three channels), the image synthesizer unit 42 which compounds two or more video signals, the image distribution apparatus 43 which distributes a synthetic video signal to a multiple channel, and two or more sound signals the voice mixing equipment 44 to mix, the voice distribution apparatus 45 which distributes a mixed sound signal to a multiple channel, the source 46 of a background image which accumulates two or more background video signals, the background sound source 47 which accumulates two or more background sound signals, and the positional information storage 48 which stores each one of positional information for every background -- and It consists of a control unit 49 which analyzes the control information from a terminal unit side, and controls each equipment. In addition, as for a video-signal bus and 440, 420 is [ a sound signal bus and 490 ] control signal buses.

[0012] First, basic actuation of this system is explained. Now, Users A, B, and C presuppose that terminal units 1, 2, and 3 are used, respectively. Moreover, each user shall picturize against the background of the wall of monochrome solid color. As a color of the wall of a background, "blue" is desirable. This is for making easy to do image composition by the below-mentioned chroma-key | *picture*

processing.

[0013] Each terminal units 1, 2, and 3 input the "virtual society" selection information as which each users' A, B, and C image and voice (it is called user image in which this was summarized) are changed into a video signal and a sound signal with image pick-up equipment 11 and a telephone transmitter 12, respectively, and a user specifies them as control information from the control information input unit 15, and transmit these video signals, a sound signal, and control information to service center equipment 4 through a communication network 5 from a transmitter-receiver 16. A sign / decryption section 161 is contained in the transmitter-receiver 16, and it compresses in order to transmit a video signal, a sound signal, and control information efficiently. *compress*

[0014] With service center equipment 4, the video signal which each terminal units 1, 2, and 3 transmit, a sound signal, and control information are received by the transmitter-receiver 41, and it decodes by each sign / decryption section 411, 412, and 413, separates into a video signal, a sound signal, and control information, and sends to the image synthesizer unit 42, voice mixing equipment 44, and a control unit 49 through the video-signal bus 420, the sound signal bus 440, and the control signal bus 460, respectively. A control unit 49 analyzes control information, chooses from the source 46 of a background image, and the background sound source 47 the background video signal and background sound signal corresponding to the "virtual society" which the user specified, respectively, and sends them to the image synthesizer unit 42 and voice mixing equipment 44. Furthermore, a control unit 49 sends the positional information of each users A, B, and C corresponding to the virtual society concerned to the image synthesizer unit 42 with reference to the positional information memory 48. In addition, the value or any value beforehand specified by the user is stored in the positional information memory 48. According to each user's positional information, the image synthesizer unit 42 compounds each user's video signal and background video signal, and sends them to the image distribution apparatus 43. Voice mixing equipment 44 mixes each user's sound signal and background sound signal, and sends them to the voice distribution apparatus 45. The image distribution apparatus 43 and the voice distribution apparatus 45 distribute the picture signal compounded, respectively and the mixed sound signal to a sign / decryption sections 411, 412, and 413. A sign / decryption sections 411, 412, and 413 encode a synthetic picture signal and a mixed sound signal, respectively, and transmit to terminal units 1, 2, and 3. *decode*

[0015] In each terminal units 1, 2, and 3, a synthetic picture signal and a mixed sound signal are received by the transmitter-receiver 16, it decodes by its sign / decryption section 161, and a synthetic video signal and a mixed sound signal are separated. Television equipment 13, projects a synthetic image and an earphone 14 outputs mixed voice. Consequently, it can talk and the users A, B, and C who are in front of each terminals 1, 2, and 3 can alternate, as met in the same location. *decode*

[0016] Next, various kinds of methods of realizing image composition are explained. Now, each users A and BC presuppose that the location is chosen so that it may stand in a line in order of A, B, and C from the left toward a background image. In this case, according to the directions from a control unit 49, the image synthesizer unit 42 compounds each user's video signal and background image by chroma-key processing, as shown in (1) of drawing 4. In this condition, if each users A, B, and C choose a silhouette image as the method of presentation of their image with the control information input unit 15, this control information will be told to a control unit 49, and will serve as a command to the image synthesizer unit 42. In case the image synthesizer unit 42 compounds each user's image and background image, it changes an applicable user's image into a silhouette image. In transform processing to a silhouette image, approaches, such as making into black level the part extracted by the chroma-key, and the approach by the fixed pattern of the source of a silhouette image prepared separately can be considered. The synthetic image corresponding to each terminals 1, 2, and 3 in the case of being based on a fixed pattern (2) of drawing 4 R> 4, (3), and (4) is shown. *decode*

[0017] Moreover, when a certain user directs to move one's location with the control information input unit 15 (for example, the migration direction, distance, etc. are directed with a mouse etc.), this is told to a control unit 49, and a control unit 49 updates the positional information of the applicable user in the positional information memory 48, and considers it as the location command to the image synthesizer

unit 42. [ in a background image ] The image synthesizer unit 42 moves to the location which had an applicable user's video signal specified according to this location command, and is compounded with other users' video signal by chroma-key processing. Moreover, if the image synthesizer unit 42 is ordered expansion of the image and visual field of the user who a control unit 42 calculates the relative-position relation between the user to whom migration was directed with reference to the positional information memory 48, and other users, and approaches at this time, migration of a certain user's image can be followed, and a specific partner's image and visual field can also be expanded and (rise image) compounded.

[0018] Next, various kinds of methods of realizing voice mixing are explained. As shown in drawing 4 , when Users A, B, and C have stood in a line in order of A, B, and C from the left now, voice mixing equipment 44 adjusts each user's vocal sound level like this arrangement. In this case, although size-related adjustment is sufficient when a voice channel is a monophonic recording, in the case of a stereo, the voice distributed to right and left is adjusted. Moreover, it is made not to return a user's voice own [ a certain ] to the user. Furthermore, when a specific partner's image and visual field are expanded, only this partner's sound signal is enlarged alternatively and other sound signals are made small. These control is the commands from a control unit 49, and is performed synchronizing with image composition.

[0019]

[Effect of the Invention] According to the virtual society implementation system of this invention, as met in the same location, the user of a large number which are present in a remote place can talk, and can alternate in a natural form. Thus, this invention system gives the new system configuration for the alternating current with people and a man.

---

[Translation done.]



\* NOTICES \*

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.\*\*\*\* shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

---

TECHNICAL FIELD

---

[Industrial Application] This invention relates to the virtual society implementation system which provides two or more human beings with the place of an alternating current of imagination with a pictorial communication technique and a voice communication technique through a communication network.

---

[Translation done.]

\* NOTICES \*

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

---

PRIOR ART

---

[Description of the Prior Art] The teleconference with voice is realized as a conventional similar system. Moreover, using the pictorial communication technique, at the so-called television conference, among two or more points, an image is sent, it suits, the image of a television machine is divided, and the method which displays two or more screens there at coincidence, and the method displayed on two or more television machines of airraid correspondence are taken.

---

[Translation done.]

**\* NOTICES \***

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

---

**EFFECT OF THE INVENTION**

---

[Effect of the Invention] According to the virtual society implementation system of this invention, as met in the same location, the user of a large number which are present in a remote place can talk, and can alternate in a natural form. Thus, this invention system gives the new system configuration for the alternating current with people and a man.

---

[Translation done.]

\* NOTICES \*

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.\*\*\*\* shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

---

TECHNICAL PROBLEM

---

[Problem(s) to be Solved by the Invention] In the above-mentioned conventional technique, an alternating current in the same location as the partner who wants to alternate mutually cannot be performed, and there is a problem that the number of points is restricted, by communication between many points.

[0004] As the purpose of this invention is whether two or more alternating human beings gathered in one location, it is to offer the virtual society implementation system with an eye on losing human being's number of limits which the alternating current which led can be performed and interchange conversation and a gesture.

---

[Translation done.]

**\* NOTICES \***

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

---

**MEANS**

---

[Means for Solving the Problem] The virtual society implementation system of this invention consists of two or more terminal units and service center equipment connected with them through the communication line. Each terminal unit The image pick-up equipment which picturizes the person image of the terminal user concerned, and the telephone transmitter which inputs this user's voice, The input unit which inputs the control information by this user, and the sending set which transmits said user's person video signal, a sound signal, and control information to center equipment, The receiving set which receives the synthetic video signal which said center equipment transmits, and a mixed sound signal, It has the television equipment which outputs said received synthetic video signal, and the telephone transmitter which outputs a mixed sound signal. Service center equipment The source of a background image which accumulates two or more background video signals, and the background sound source which accumulates two or more background sound signals, The receiving set which receives the person video signal, the sound signal, and control information which said each terminal unit transmits, The image synthesizer unit which compounds the person video signal from said each terminal unit, and the background video signal chosen from said source of a background video signal, The voice mixing equipment which mixes the sound signal from said each terminal unit, and the background sound signal chosen from said background sound source, The distribution apparatus which distributes said synthetic video signal and mixed sound signal to each terminal unit, The control information from said each terminal unit is analyzed, and it is characterized by having the control unit which performs image composition of the selection control of the background video signal from said source of a background image, and a background sound source, or a background sound signal, said image synthesizer unit, and voice mixing equipment, and control of voice mixing.

---

[Translation done.]

\* NOTICES \*

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

---

OPERATION

---

[Function] A user's image and voice which are in front of each terminal unit are changed into a video signal and a sound signal by image pick-up equipment and the telephone transmitter, and are sent out to a communication network by them, respectively. Each of this user's image and sound signal are sent to service center equipment, are compounded with a background image and a background sound there, and are \*\*\*\* about the imagination world "virtual society" -- \*\*. Through a communication network, this image and voice of "virtual society" that were compounded are sent to each user's terminal unit, and are reproduced by television equipment and the earphone.

---

[Translation done.]

\* NOTICES \*

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.\*\*\*\* shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

---

DESCRIPTION OF DRAWINGS

---

[Brief Description of the Drawings]

[Drawing 1] It is the block diagram of the whole virtual society implementation system of this invention.

[Drawing 2] It is drawing showing the example of a configuration of each terminal unit of drawing 1 .

[Drawing 3] It is drawing showing the example of a configuration of the service center equipment of drawing 1 .

[Drawing 4] It is drawing showing an example of the synthetic image displayed on each terminal unit.

[Description of Notations]

1-3 Terminal unit

4 Service Center Equipment

5 Communication Network

11 Shooting Equipment

12 Telephone Transmitter

13 Television Equipment

14 Earphone

15 Control Information Input Unit

16 Transmitter-receiver

41 Transmitter-receiver

42 Image Synthesizer Unit

43 Image Distribution Apparatus

44 Voice Mixing Equipment

45 Voice Distribution Apparatus

46 Source of Background Image

47 Background Sound Source

48 Positional Information Memory

49 Control Unit

---

[Translation done.]

\* NOTICES \*

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

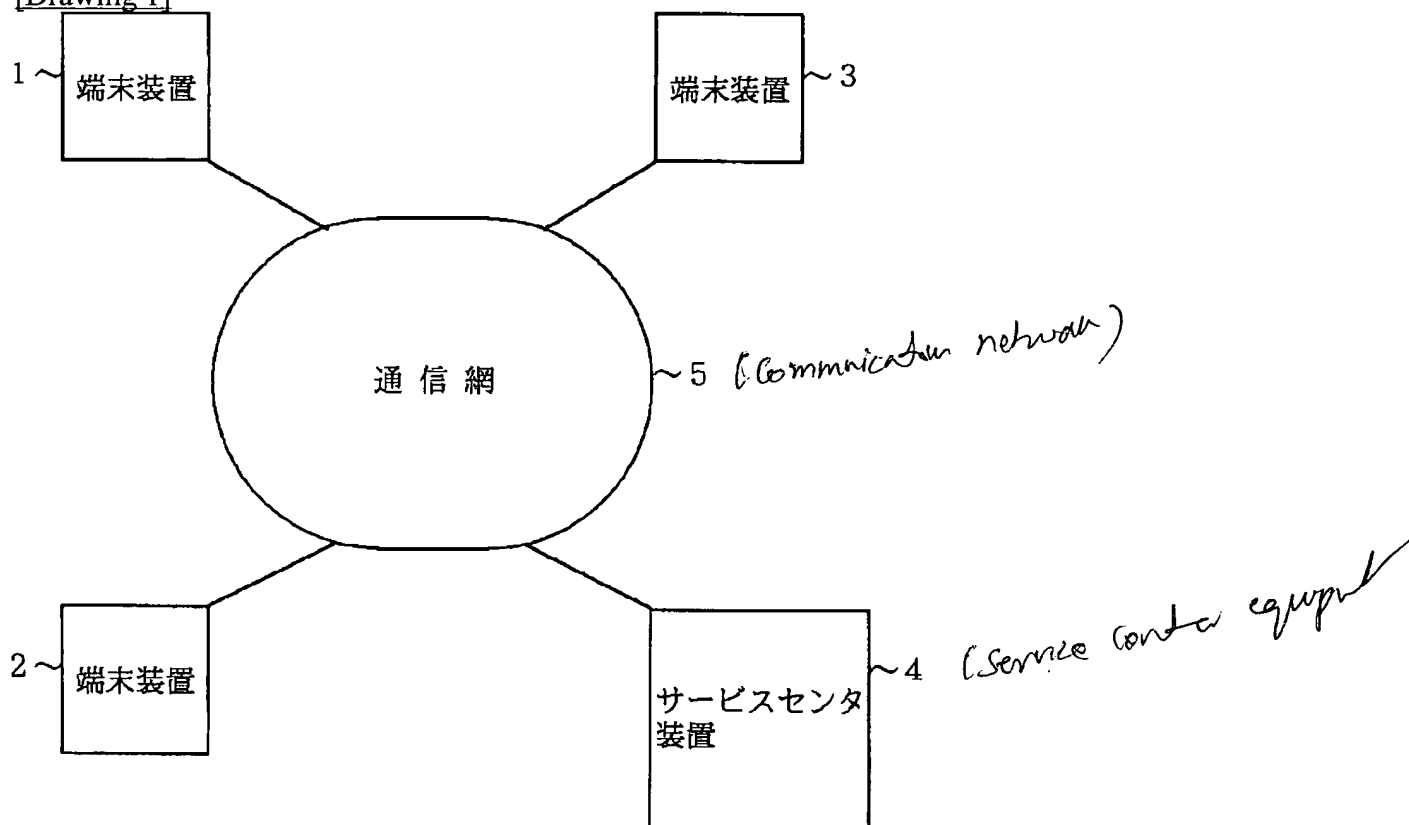
1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

---

DRAWINGS

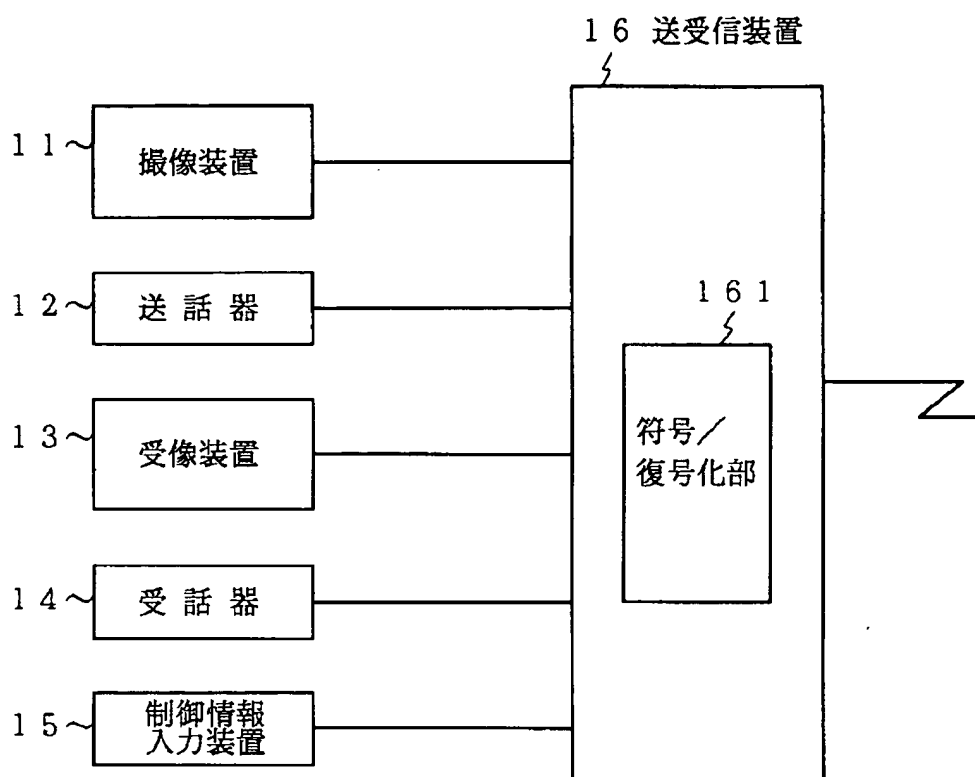
---

[Drawing 1]



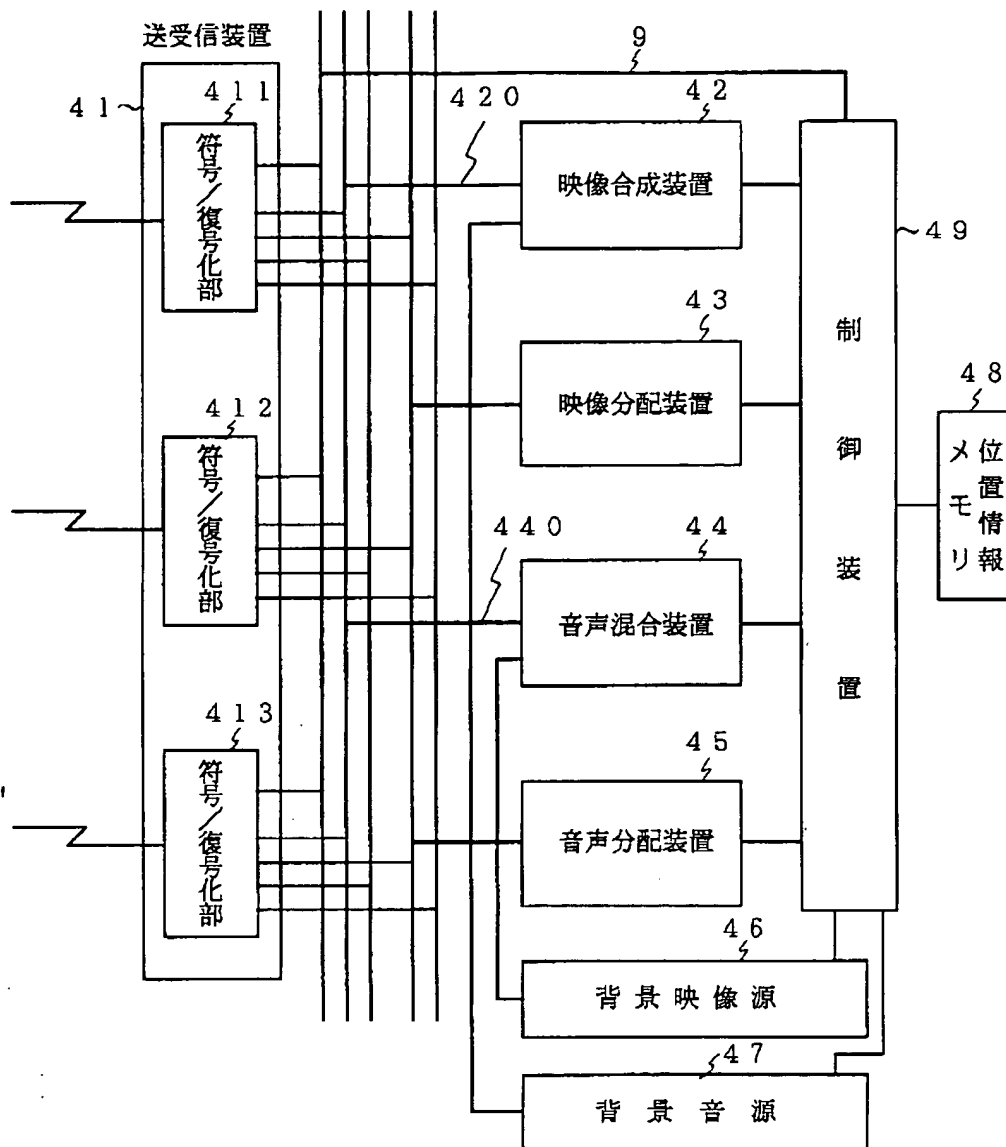
[Drawing 2]





[Drawing 3]

15: Control information input unit  
 14: Earphone  
 13: Television equipment  
 12: Telephone transmitter  
 11: image pickup equipment  
 16: transmitter/receiver



[Drawing 4]

Service Center equipment

41: Transmitter/receiver

42: Image synthesizer

43: Image distribution apparatus

44: Voice mixing equipment

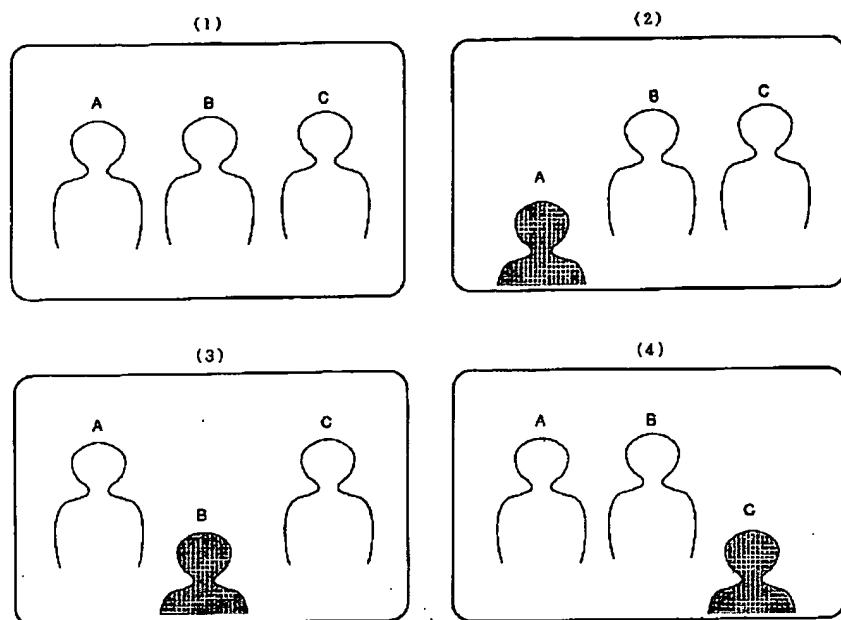
45: Voice distribution equipment

46: Source of background image

47: background sound source

48: position information storage unit

49: ~~49~~ Control unit



---

[Translation done.]